DOCUMENT RESUME

ED 053 793 24 PS 004 812

AUTHOR	Scott, Joseph A.; Frayer, Dorothy A.
TITLE	Learning by Discovery: A Review of the Research Methodology. Report from the Project on Variables and Processes in Cognitive Learning.
INSTITUTION	Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
REPORT NO	UW-WRDCCL-WP-64
BUREAU NO	BR-5-0216
PUB DATE	Dec 70
CONTRACT	OEC-5-10-154
NOTE	35p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	*Discovery Learning, *Learning, *Learning Processes, *Research Methodology, Retention, *Rote Learning, Transfer of Training

ABSTRACT

Research comparing discovery and expository methods of presentation has yielded conflicting results. A review of the research on discovery learning is provided in this paper, focusing on the methodology of each study. Conclusions are drawn concerning the effects of discovery methods of presentation on initial learning, transfer, and retention. Use of a standardized concept learning task is recommended for future research on discovery methods of presentation. (Author)

PA-24
BR-5-0216

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED053793

Working Paper No. 64

LEARNING BY DISCOVERY: A REVIEW OF THE RESEARCH METHODOLOGY

By Joseph A. Scott and Dorothy A. Frayer

Report from the Project on
Variables and Processes in Cognitive Learning

Principal Investigators: Herbert J. Klausmeier, Robert E. Davidson,
Joel R. Levin, Thomas A. Romberg, B. Robert Tabachnick,
Alan Voelker, Larry Wilder, Peter Wolff
Technical Development Program Director: Mary R. Quilling
Research Associate: Dorothy A. Frayer

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

December 1970

PS 004812

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154

NATIONAL EVALUATION COMMITTEE

Samuel Brownell
Professor of Urban Education
Graduate School
Yale University

Launor F. Carter
Senior Vice President on Technology and Development
System Development Corporation

Francis S. Chase
Professor
Department of Education
University of Chicago

Henry Chauncey
President
Educational Testing Service

Martin Deutsch
Director, Institute for Developmental Studies
New York Medical College

Jack Edling
Director, Teaching Research Division
Oregon State System of Higher Education

Elizabeth Koontz
Wage and Labor Standards Administration, U.S. Department of Labor, Washington

Roderick McPhee
President
Punahou School, Honolulu

G. Wesley Sowards
Director, Elementary Education
Florida State University

Patrick Suppes
Professor
Department of Mathematics
Stanford University

***Benton J. Underwood**
Professor
Department of Psychology
Northwestern University

RESEARCH AND DEVELOPMENT CENTER POLICY REVIEW BOARD

Leonard Berkowitz
Chairman
Department of Psychology

Archie A. Buchmiller
Deputy State Superintendent
Department of Public Instruction

Robert E. Grinder
Chairman
Department of Educational Psychology

Russell J. Hosler
Professor, Curriculum and Instruction

Clauston Jenkins
Assistant Director
Coordinating Committee for Higher Education

Herbert J. Klausmeier
Director, R & D Center
Professor of Educational Psychology

Stephen C. Kleene
Dean, College of Letters and Science

Donald J. McCarty
Dean
School of Education

Ira Sharkansky
Associate Professor of Political Science

B. Robert Tabachnick
Chairman, Department of Curriculum and Instruction

Henry C. Weinlick
Executive Secretary
Wisconsin Education Association

M. Crawford Young
Associate Dean
The Graduate School

EXECUTIVE COMMITTEE

Edgar F. Borgatta
Brittingham Professor of Sociology

Anne E. Buchanan
Project Specialist
R & D Center

Robin S. Chapman
Research Associate
R & D Center

Robert E. Davidson
Assistant Professor, Educational Psychology

Frank H. Farley
Associate Professor, Educational Psychology

Russell J. Hosler
Professor of Curriculum and Instruction and of Business

***Herbert J. Klausmeier**
Director, R & D Center
Professor of Educational Psychology

Wayne Otto
Professor of Curriculum and Instruction (Reading)

Robert G. Petzold
Associate Dean of the School of Education
Professor of Curriculum and Instruction and of Music

FACULTY OF PRINCIPAL INVESTIGATORS

Vernon L. Allen
Professor of Psychology

Ted Czajkowski
Assistant Professor of Curriculum and Instruction

Robert E. Davidson
Assistant Professor of Educational Psychology

Gary A. Davis
Associate Professor of Educational Psychology

M. Vere DeVault
Professor of Curriculum and Instruction (Mathematics)

Frank H. Farley
Associate Professor of Educational Psychology

Lester S. Golub
Lecturer in Curriculum and Instruction and in English

John G. Harvey
Associate Professor of Mathematics and of Curriculum and Instruction

Herbert J. Klausmeier
Director, R & D Center
Professor of Educational Psychology

Donald Lange
Assistant Professor of Curriculum and Instruction

James Moser
Assistant Professor of Mathematics Education; Visiting Scholar

Wayne Otto
Professor of Curriculum and Instruction (Reading)

Milton O. Pella
Professor of Curriculum and Instruction (Science)

Thomas A. Romberg
Associate Director, R & D Center
Professor of Mathematics and of Curriculum and Instruction

B. Robert Tabachnick
Chairman, Department of Curriculum and Instruction

Richard L. Venezky
Assistant Professor of English and of Computer Sciences

Alan Voelker
Assistant Professor of Curriculum and Instruction

Larry Wilder
Assistant Professor of Curriculum and Instruction

Peter Wolff
Assistant Professor of Educational Psychology

MANAGEMENT COUNCIL

Herbert J. Klausmeier
Director, R & D Center
V.A.C. Henmon Professor of Educational Psychology

Mary R. Quilling
Director
Technical Development Program

Thomas A. Romberg
Associate Director

James Walter
Director
Dissemination Program

Dan G. Woolpert
Director
Operations and Business

* COMMITTEE CHAIRMAN

STATEMENT OF FOCUS

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This Working Paper is from the Project on Variables and Processes in Cognitive Learning in Program 1. General objectives of the Program are to generate knowledge about concept learning and cognitive skills, to synthesize existing knowledge and develop general taxonomies, models, or theories of cognitive learning, and to utilize the knowledge in the development of curriculum materials and procedures. Contributing to these Program objectives, the Cognitive Learning Project has these objectives: to ascertain the important variables in cognitive learning and to apply relevant knowledge to the development of instructional materials and to the programming of instruction for individual students; to clarify the basic processes and abilities involved in concept learning, and to develop a system of individually guided motivation for use in the elementary school.

TABLE OF CONTENTS

		Page
I.	Introduction	1
	Background of the Problem.	1
	Problems in the Discovery Research	2
	Semantic inconsistencies	2
	Conceptual issues.	3
	What is Discovered?.	4
	Three Dimensions of Discovery Learning	5
	The Case for a Standardized Task	6
	Purposes of the Paper.	7
II.	Research on Learning by Discovery.	8
	Discovery vs. Non-Discovery.	9
	Amount of Guidance	11
	Effects of Verbalization on Discovery.	19
	Summary.	19
III.	Summary and Implications	21
	Summary.	21
	Implications for Future Studies.	22
	A Standardized Task for Discovery Research	23
	Measurement of the Effects of Discovery Learning	26
	References	28

I

INTRODUCTION

Background of the Problem

Dewey (1963) stated that "the history of educational theory is marked by opposition between the idea that education is developed from within and that it is formation from without" (p. 17). Reference to this conflict can be found in philosophical writings for the past two thousand years. Each generation which has debated the question has introduced its own assumptions, philosophies,and criteria toward its resolution. The present generation, likewise, has examined the problem from its own viewpoint.

Psychologists have reformulated the question to ask "Which is superior, discovery or expository learning?" This question is amenable to examination in laboratory settings. Research of this nature has been in progress for over four decades. The results to date, however, have been equivocal. Wittrock (1966) prefaced his review of the research on discovery with the warning that the current state of the research on discovery is very disappointing and precludes any important conclusions about teaching and learning. Summarizing the findings of the Conference of Learning by Discovery, Morrisett (1966, p. 179) concluded that "research on the topic of discovery . . . is relatively

impoverished . . . first in the range of variables which have been considered . . . [and second] in the subject matters that have been studied."

The conclusion of the Conference on Learning by Discovery was that even such a basic consideration as the definition of discovery was not agreed on by the participants. Howard Kendler (1966) in his concluding remarks at the conference pointed to the confusion and lack of communication involved in the use of the word "discovery" and called for its abandonment. No one has, however, called for abandonment of research on the basic question, which is the assessment of the effects on learning and retention of various methods of presenting stimulus materials to students in classroom situations. And it appears unlikely that the word "discovery" itself will be abandoned, so the effort towards a more universally acceptable definition must continue.

Problems in the Discovery Research

Wittrock (1966) classified the many problems associated with the research on discovery learning into three categories: semantic inconsistencies, conceptual issues, and methodological problems. This paper will focus on methodological problems but it is well to realize that semantic inconsistencies and conceptual issues permeate all of the studies which will be examined and it might be well to outline here Wittrock's characterization of these problems.

Semantic inconsistencies. The first semantic problem noted by Wittrock (1966) is the lack of operational definitions and objective indices for the term "discovery." The clearest evidence of this problem

was the conference itself and the obvious confusion and lack of agreement among the participants as to what they meant by "discovery." Robert Glaser (1966) stated that discovery sequences are characterized by two properties: inductive sequence and trial and error learning. Gagné (1966) said that discovery involves inferring an internal process of search and an internal process of selection, and suggested that discovery could occur at each of the seven levels in his hierarchy of learning. Kagan (1966) equated the "inferential approach" with the discovery approach and stated that the discovery method requires that the child infer a major principle "without excessive guidance" from an external source. Bruner's (1961) definition of discovery was so broad as to "include all forms of obtaining knowledge for oneself by the use of one's own mind" (p. 22). Later in the same article, however, he narrowed his definition of discovery to "rearranging or transforming evidence in such a way that one is enabled to go beyond the evidence so reassembled to additional new insights" (p. 22). From these examples, it is obvious that scholars in the area have not reached agreement.

A second semantic problem noted by Wittrock (1966) is that treatments are labeled in terms of responses they are said to produce. Wittrock objects to labeling treatments with words such as "rote" and "discovery" since they denote internal processes rather than stimuli.

Conceptual issues. Wittrock raised several issues related to the conceptualization of discovery. The most cogent of these issues are: (a) What criteria should be used? The variety of

dependent variables makes comparison of studies a difficult enterprise. (b) Should discovery be considered merely a way to learn subject matter, or should it be an end in itself? Bruner holds that learning to discover is at least as important as learning by discovery. This question cannot be resolved by research, however. (c) What is discovered or discoverable? This question remains unanswered and perhaps cannot be answered by research.

These and several other questions and problems of this nature plague the literature on discovery learning and render drawing any kind of conclusions regarding teaching and learning exceedingly difficult.

What is Discovered?

The question of the nature of what is discovered is epistemological rather than psychological but some assumptions about what is potentially discoverable must be made before stimulus materials are prepared and data collected. By examining the tasks used in the studies reported in Chapter 2 of this paper it will be noted that what is usually taken to be discoverable is what Gagné (1965) would designate as a principle. Most studies required subjects to find a rule or a heuristic of a complex nature.

For example, Katona (1940) used a task in which Ss had to discover a rule for solving the "card-trick" problem. Ewert and Lambert (1932) used a task in which Ss had to discover the general principle which would allow them to complete the puzzle, involving the transfer of discs among three circles, in the fewest possible moves. Craig (1953) used

a task which consisted in the elimination of a word which "did not belong" from a group; Subjects (Ss) had to discover the grouping principle. In each of these examples, what Ss had to discover could be described as a principle. Likewise, what is taken to be discoverable in almost all studies reported in Chapter 2 can be classified as a principle à la Gagné.

This raises two problems. Since little research has been done on principle learning *per se*, it is difficult to specify the dimensions of these principles so that the relative difficulty can be compared across tasks. Beyond this, it is difficult to control extraneous variables when conducting discovery research, since little information is available on variables which affect the learning of principles. A considerable body of data, however, has been accumulated on concept learning over the past decades, and the nature of the variables which affect concept learning is well understood. Research on discovery of concepts might, therefore, yield more consistent results than research on discovery of principles.

Three Dimensions of Discovery Learning

Studies of discovery learning can be subdivided into three groups: studies which have juxtaposed expository and discovery learning (e.g., Swenson, 1949; Thiele, 1938); studies which have examined effects of amount of guidance (e.g., Craig, 1956; Wittrock, 1963); and studies of the effects of verbalization on discovery learning (e.g., Hendrix, 1947; Gagné & Smith, 1962). A fourth group of studies are those which have attempted to answer questions regarding the motivational effects of discovery *versus* expository learning (e.g., Kersh, 1958, 1962). In these studies,

PS 004812

however, motivational effects for the most part have not been isolated from effects of amount of guidance and will consequently be reported with studies of effects of amount of guidance.

The question might be raised as to the viability of a concept learning task, as proposed in the last section, for researching these dimensions. It will be pointed out in the next chapter that the lack of specifiable dimensions of the tasks and stimulus materials used in previous studies makes comparison of the studies or resolution of the contradictory findings extremely difficult. This is especially true of studies examining the effects of amount of guidance. Much of this difficulty could be eliminated by the use of a concept learning task. Concept learning tasks could be compared in terms of number of attributes of the concept to be discovered, sequence and number of examples presented, ratio of positive to negative instances, cues provided per instance, etc. Findings would, therefore, be much more informative than those of present studies where non-specific labels such as "intermediate amount of guidance" are used. Similar benefits in comparability of findings of various studies would also accrue from the use of a concept learning task in the investigation of general effects of expository vs. discovery learning and of the effects of verbalization on discovery learning.

The Case for a Standardized Task

Bruner, Goodnow, and Austin (1956) described a set of stimulus materials and tasks which became a standard for most of the research on concept learning during the next 15 years. Bruner's paradigm and the materials and tasks he utilized had many advantages. They were describable in exact

terms. Both what was varied and what was controlled could be exactly specified. Widespread use of the paradigm resulted in a large body of comparable data in a series of replicable studies, and in findings which were, therefore, verifiable. The area of discovery learning lacks such a well-specified task. It will be one of the functions of this paper to specify what such a task might be.

Purposes of the Paper

This paper has three purposes:

1. To provide a representative review of the literature on discovery learning. In this review, the focus will be on the methodology of each study, with particular emphasis on independent and dependent variables, questions asked by the study, and legitimacy of the findings.
2. To provide a summary of the findings and conclusions that may be legitimately drawn from the research reviewed.
3. To describe a standardized task that may be used in future studies of discovery learning.

II

RESEARCH ON LEARNING BY DISCOVERY

Research on the effects of learning by discovery has been conducted over the past four decades. The following review will not attempt to be exhaustive since many reviews of the research have appeared, the latest being that of Wittrock (1966). This review will concentrate instead on a representative selection of studies which will illustrate the kinds of questions asked and the methodology of the studies which sought to answer these questions.

It is possible to divide the experiments performed into three groups on the basis of the independent variables examined: discovery vs. non-discovery methods of presentation; amount of guidance; and verbalization or no verbalization. Most of the studies reviewed measured either retention or transfer or both. Exceptions to this will be pointed out in context. For most of the studies reported here the independent and dependent variables will be noted. In addition, the tasks employed, type of subject, duration of learning, and retention intervals will be recorded for whatever light they may shed on the results obtained.

Discovery vs. Non-Discovery

A number of studies have examined the question of whether discovery methods of presenting material produce higher scores on tests of retention and transfer than expository or rote learning methods. In one of the earliest experiments of this kind, McConnell (1934) used Second Grade Ss in a 7-month study. The stimulus materials were 100 addition and 100 subtraction facts. The methods were: Authoritative, in which Ss were told to memorize the facts, and Discovery, in which Ss were told to discover the generalization involved in the task. The results showed the Authoritative method to be best on speeded retention, but discovery to be superior on transfer tests.

Thiele's (1938) design and methodology was similar to McConnell's. Thiele's task also consisted in learning 100 addition facts and the subjects were again Second Graders. Treatments were Generalization, in which Ss were told to look for a generalization, and Drill in which facts were presented without any attempt by E to relate them to each other. Training time was 7 weeks. Performance of the Generalization group was superior on all measures of retention and transfer.

Katona (1940) used a somewhat different task: Ss, graduate students in psychology, were assigned to one of three groups. Ss in the Memorization Group were told the correct sequence of cards required to perform a trick; Ss in the Understanding Group were told the way in which the solution could be derived; Ss in the Control Group received no training. Training time was 4 minutes. Results showed the Memorization Group to be superior on a test of immediate retention, while the Understanding

Group was superior on a test of transfer to similar problems. Both groups were superior to the Control Group. The results were replicated in a second experiment, with one additional result. On a retest 4 weeks after training, the Understanding Group was superior both on memory and transfer.

Swenson's (1949) Ss were Second Grade students. The task was learning 100 addition facts and the duration of training was 16-1/2 weeks, with lessons of 25 minutes each day. Treatments were: Generalization Method, encouraging the students to build up inter-relationships; Drill Method, presenting facts in miscellaneous order; and Drill-Plus Method, duplicating standard teaching procedures. This included manipulation of objects in addition to drill on facts. Results showed little difference among groups on initial learning, but the Generalization group was higher on both retention and transfer.

Anderson (1949) used Fourth Grade students and applied his treatments to the material taught in the regular math curriculum from November through May. Treatments were Drill and Meaning methods, essentially the same as the Drill and Generalization methods described in the Swenson study. On standardized arithmetic tests he found no differences among groups as a function of method of teaching. Since the tests were administered on the materials presented during the year this can be considered a test of retention. A test of mathematical thinking, which can be considered a transfer test, was also given. Ss of high ability but inferior achievement, who had received the Meaning method of training, performed best on this test.

The general, though by no means universal, finding of these studies appears to be that discovery methods are not superior to rote or drill methods when the criterion is immediate learning or short-term retention but become superior when the criterion is either long-term retention or transfer. The findings of these studies do not appear to be related to task, age of Ss, or length of training time.

It is possible to speculate on the sources of effects due to presentation method. In the drill methods, facts were presented to Ss. Their task was to remember these facts. The task for Ss in the generalization groups was to find a generalization. Since the task differed between groups they may have held different expectations about the test and would be prepared to answer different questions. Learning of a generalization should produce better performance on a transfer test than merely learning facts. No experiment was performed to measure differences among Ss who discovered a generalization, Ss who were given a generalization, and Ss who learned facts alone. Such a comparison might suggest how much of the effect was due to Ss discovering a generalization as opposed to Ss using a generalization.

Amount of Guidance

By far the largest group of studies dealing with discovery have examined this independent variable. Ewert and Lambert (1932) varied amount of guidance using four treatments. Method 1 gave Ss the objectives of the problems and rules of procedure. Method 2 also gave Ss objectives and rules and, in addition, asked Ss to find one general principle applicable to all problems. Method 3 gave Ss objectives, rules, and

general principles. Method 4 added a demonstration to the procedures of Method 3. The task was a puzzle requiring the movement of graduated discs among three circles (similar to the classic "Towers of Hanoi" problem). Results showed greater guidance to be most effective; the methods, in the order of increasing superiority, were 1, 2, 4, 3. The dependent variables were time to criterion and number of moves.

Stacey (1949) used Sixth Grade Ss to compare the effectiveness of five methods, two of which could be characterized as discovery methods and three as authoritative methods. Amount of guidance varied among the five methods. The task consisted of elimination of one word that "did not belong" from a set of five words. Although the conclusions drawn favored discovery methods and minimum guidance, few of the findings reached a .05 level of significance.

Craig (1953) used a task similar to Stacey's. Ss were males, recent college graduates who were being commissioned as second lieutenants in the U. S. Air Force. Four levels of guidance were used: zero clues, (Group Z); grouping of stimulus material to maximize discovery of relationships (Group G); information that the stimulus material was grouped according to some principle, (Group GX); and a fourth treatment where Ss, in addition to the receiving the information given to Group GX, were told the grouping principle (Group GXP). Results showed that number of errors to solution of the problem was inversely related to amount of guidance. Amount of transfer increased in direct relationship to the amount of guidance. In a second study, Craig (1956) used college Ss. The task was similar to that used in the previous study. The treatments

consisted of either no clues (Independent Group) or a short general statement of the relationship among the items (Directed Group). The Directed Group was superior on a test of initial learning and on a retention test administered 31 days later, but not on a retention test administered 3 and 17 days after completion of the task. No differences were found between the groups on a transfer test.

Kittel (1957) also utilized a task similar to Stacey's and Craig's but with Sixth Grade students. Three levels of guidance were labeled Minimum, Intermediate, and Maximum. The Intermediate Direction group used in this study appears to be equivalent to the GXP (maximum direction) group in the Craig (1953) study and the Directed Group in the Craig (1956) study. Results showed the Intermediate Direction group superior on all measures of retention and transfer.

Corman (1957) also varied amount of guidance using the Katona matchstick task (Katona, 1940). Twelfth Grade students were Ss in this study. The results indicated a complex set of interactions among mental ability, kind of information given, and dependent measures.

Haslerud and Meyers (1958) used a code deciphering and enciphering task. Ss were college students. Treatments were No Directions regarding the code and Specific Directions regarding the code. The No Direction group was found to be superior on a transfer task. The validity of the analysis has, however, been disputed many times, e.g., Wittrock (1966), Cronbach (1966).

Wittrock (1963) also examined the effects of varying amounts of guidance on a discovery task. Using college Ss and a code deciphering task, he formed four groups, varying in degree of guidance: Rule

given, Answer given (RgAg); Rule given, Answer not given (RgAng); Rule not given, Answer given (RngAg); and Rule not given, Answer not given (RngAng). On initial learning, the RgAg and RgAng groups were significantly better than the other two groups, but did not differ from one another. On a combined test of retention and transfer, RgAng was significantly better than RgAg and RngAng. Wittrock interpreted these findings as evidence that maximal guidance produces greatest initial learning, while an intermediate amount of guidance produces greatest retention and transfer.

Forgus and Schwartz (1957) attempted to test the validity of earlier Katona studies using a new 26-symbol alphabet. Female college students served as Ss in this study. Three treatments were used. In Treatment O, Ss were told the principle underlying construction of the alphabet; in Treatment P, Ss were told there was a principle and were asked to describe it; and in Treatment M, Ss were asked to memorize the alphabet. Both P and O groups, though not significantly different from each other, were significantly higher than Group M on both retention and transfer tests one week after training.

Kersh (1958, 1962) has conducted two studies varying the amount of guidance. In the 1958 study, Kersh compared the effects of three treatments: No Help, Direct Reference (Ss were given perceptual aids), and Rule Given. The task consisted of verbalizing mathematical generalizations and applying them to new problems. Ss were college students. The dependent measures were retention, transfer, and the heuristic which Ss used to solve problems on a retest. The data failed to support the hypothesis that the Direct

Reference Group would be superior on a test of retention and transfer. An additional finding of interest was that among 13 Ss in the No Help Group who failed to discover the rule during the learning period, 10 used acceptable methods when retested 4 weeks later. This is in contrast to the other treatment groups in which there was a decrease in use of acceptable methods from test to retest. Kersh attributed this to a differential motivational level as a function of treatments and concluded that motivation was a more important factor than understanding in the effects of discovery learning. Kersh's second study (1962) was designed to examine this motivational effect. The task was the same as in the previous study but Ss were high school students. Three treatments were used: Directed Learning, Guided Discovery, and Rote Learning. Ss in the Directed Learning Group were given programmed booklets which presented the mathematical rules and explanations of the rules. Ss in the Guided Discovery Group discovered the explanations under a Socratic method of teaching and were presumed to be most highly motivated. Ss in the Rote Learning Group were apparently not given an explanation of the rules. Results showed the Rote Learning roup to be superior on all measures to the other groups. The Guided Discovery Group was superior to the Directed Learning group on all measures. Kersh's characterization of the Rote Learning treatment is rather difficult to accept, however, and it is the authors' opinion that the Rote Learning Group may, in fact, have been a "pure" discovery group.

Gagné and Brown (1961) assigned Ninth and Tenth Grade boys to three experimental groups: Rule and Example (R & E), Discovery (D), and Guided

Discovery (GD). The task consisted of learning to sum several number series. The GD Group performed significantly better than the D Group and the R & E Group; the D Group, however, performed significantly better than the R & E Group.

Grote (1960) used Eighth Grade Ss. The task consisted of learning the principles of simple machines. The independent variables were method of presentation and sequence of presentation methods. Two methods of presentation were used: Direct-Detailed in which S was told all E wanted him to know (an expository method); and Directed Discovery in which a minimum of basic information was given, followed by leading questions. Each S learned two principles and the methods of presenting the two principles varied systematically so that there were four conditions: Direct-Detailed, followed by Direct-Detailed; Direct-Detailed followed by Directed Discovery; Directed Discovery followed by Direct-Detailed; and Directed Discovery followed by Directed Discovery. A test of immediate acquisition after completion of the lessons on the first principle showed the Direct-Detailed method to be superior. Subsequent tests of retention and transfer showed several unexplainable interactions between sequence of method of presentation and retention interval, but no identifiable pattern.

Moss (1960) used the same independent variables: Direct-Detailed and Directed Discovery. Ss were junior and senior high school students. The task consisted of learning the technique of letterpress imposition. This task requires the arrangement of pages of type to be printed on one side of a sheet so that when cut, folded, and trimmed, they will fall in numerical sequence. Six measures were obtained: initial learning, retention after 1 week, retention after 6 weeks, immediate transfer,

transfer after 1 week, and transfer after 6 weeks. There were no significant differences between the presentation methods on any of these measures.

Tomlinson (1962) found some intriguing, if unexplainable, interactions between presentation method and type of test and between method of presentation and grade of Ss. Tomlinson used Junior and Senior High School students and the task consisted of learning information concerning the metallurgy of carbon steel. Four methods of presentation were used: an Inductive Method (I) in which Ss were given examples followed by generalizations; a Deductive Method (D) in which Ss were given generalizations first, followed by examples; an Inductive-Discovery Method (ID) in which Ss were given examples, generalizations and questions; and an Inductive-Discovery-Confirmation method (IDC), in which Ss were given examples, generalizations, questions, and answers. Results were complex and difficult to interpret. On a test of immediate acquisition, using a true/false test, the I and ID Groups were significantly better than the D group, though not different from each other. When a multiple-choice test was used, however, the I Group was significantly better than all others and the D group was significantly better than the ID group. On retention and transfer tests administered 1 week after completion of the lessons, no differences were found as a function of method of presentation. An interaction between class and presentation method was, however, noted. The ID method was better for Junior High School Ss and the IDC method was better for Junior High School Ss. On a test administered 5 weeks later the IDC Group was found to be significantly inferior to all others and the I Group significantly better than the ID Group. No interaction with ability level was noted.

The authors find it impossible to reconcile the results of the studies just cited. One may find some evidence for the superiority of none, some, or much guidance. Although more evidence appears to indicate that some guidance produces better retention and transfer, it would appear on closer inspection that kind rather than quantity of guidance is most important. This is suggested very strongly by five of the studies just cited. The superior group in the Ewert and Lambert (1932) study had been given the objectives, rules, and general principles. The superior group in the Craig (1953) and Kittel (1957) studies had been given the grouping principle. In the Craig (1956) study the Ss who scored highest had been given a statement of the relationships among the items. Of the two groups which performed best in the Forgus and Schwartz (1957) study, one was told the principle underlying the construction of the new alphabet, the other was told to find the underlying principle. From these studies then it would appear that giving a rule or principle has a strong effect and that the important variable is the kind of information, rather than the amount of guidance. This suggests the necessity for a more detailed description of the task, stimulus material, and kind and amount of guidance in preparing and reporting studies. A experiment might be suggested in which kind and amount of guidance are separated to assess the relative contribution of each. It is likely, too, that the type of task (e.g., identifying a concept as opposed to learning a skill) will also interact with both kind and amount of guidance.

Effects of Verbalization on Discovery

Only three studies to date have examined the effects of verbalization on discovery. Hendrix (1947) used High School and college Ss and three treatments. In Method 1 Ss were told the principle. In Method 2 Ss were given a series of problems leading to discovery of the principle, but were not asked to verbalize it. Method 3 required Ss to verbalize the principle. Differences among the groups were not significant, but appeared to favor the non-verbalizing group. Schwartz (1948) examined the importance of verbalization in concept formation and found that most Ss who learned the concept and who could transfer to another sorting concept using the same principle were unable to verbalize the principle they were using; that unsuccessful attempts at verbalization negatively affected performance on subsequent tasks; and that those who could verbalize the principle were successful on subsequent tasks. Gagné and Smith (1962) used the three-circle problem used by Ewert and Lambert (1932) and examined the effects of verbalization and solution set (instruction to search for principle). Results showed no effects of solution set but the verbalization groups were significantly superior on number of moves and time to criterion. Not enough data are yet at hand to warrant conclusions about the effect of verbalization on discovery learning.

Summary

The literature to date has shown that discovery learning is superior to rote learning on measures of long-term retention and transfer but not superior on measures of immediate acquisiton. When the dimensions of

amount of guidance and verbalization are introduced, however, complex interactions between stimuli and as yet undetermined variables appear to occur such that definitive statements are unwarranted at this time.

GPO 822-858-4

III

SUMMARY AND IMPLICATIONS

Summary

Two findings emerge from the studies cited in the previous chapter. It appears that method of presentation does differentially affect performance on tests of immediate acquisition, retention, and transfer. The general finding is that discovery is not better than rote or expository methods of presentation when the dependent variable is immediate acquisition of the material presented. Discovery usually produces better scores on retention and transfer tests than rote or expository methods.

The second finding of some consequence is that amount of guidance has a significant effect on performance on tests of immediate acquisition, retention and transfer. Most studies have found that an intermediate amount of guidance produces best performance. The authors feel, however, that the important variable in these studies was not the amount of guidance but the kind of information given. This assertion can be empirically tested.

There are enough indicators of the positive effects of discovery methods of presenting stimulus materials to give encouragement to those who advocate the use of discovery methods in the classroom. For psychologists and others who wish to understand or explain,

rather than use the results of the research to date, much disentangling of variables remains.

Implications for Future Studies

Attributing effects to method of presentation when the method of presentation is described by such broad terms as "discovery," "generalization," "inductive," and "rote learning" must be dissatisfying to any behavioral scientist. The logical question which follows is "What is it about this presentation method that affects the results?" An answer such as "Why, it's because the child discovered the answer himself; that's why he remembers better," will not suffice. The answer is non-explanatory. Several clues are offered by the summarizing remarks in this and the previous chapter as to the way in which one might proceed. For instance, the early studies suggested the superiority of the Generalization method. The authors have noted the absence of a study to test whether discovery of a generalization is necessary or whether simply having one is sufficient to produce the superior results. Results of the experiment on amount of guidance would suggest that the latter is a very important variable. The relative effects of kind and amount of guidance need to be assessed.

Beyond this continued investigation of details, a more fundamental examination of the entire question is needed. A number of writers from Rousseau to Bruner have suggested that the motivational effects of learning by discovery are of a positive nature. With the exception of Kersh's studies, which are certainly not conclusive, no one has seriously investigated this variable. It is generally ignored and

is maintained as a confounding factor in most studies. Other important variables which have remained largely unexamined and uncontrolled are the kind of task and its interaction with the total stimulus situation, the history of the Ss, and the measuring instruments.

A Standardized Task for Discovery Research

In the first chapter of this paper the need for a standardized task was noted. At this point, having reviewed a number of studies representative of the literature in the field, it is possible to specify further what the nature of that task might be. It was noted that in many studies the subjects in the various experimental groups were given different tasks. For example, in the Swenson (1949) study, Ss in the Drill group were required to remember a series of addition facts: Ss in the Generalization Group were required to find a generalization. At the end of the experimental treatment the two groups had learned two different kinds of responses. It is hardly surprising that they should perform differently on subsequent tests. It is very difficult to justify attributing performance differences in such studies solely to method of presenting the materials. Differences may be wholly or in great part due to the differences in the content of what the child has learned, not in how he learned it.

For a meaningful comparison between methods of presentation, Ss in all experimental conditions must have the same task. For example, suppose the concept quadrilateral is presented. The relevant attribute of quadrilateral is "four sides and only four sides." The task for the discovery group should be to discover that a figure

called a quadrilateral has four and only four sides and Ss in that group should have this information by the end of the lesson. The task for the expository group should be to learn that a figure called a quadrilateral has four and only four sides, and Ss in that group should have this information by the end of the lesson. What is necessary, then, is that both kind and degree of original learning must be equated at the end of the lesson or presentation. If this is so then subsequent differences in forgetting rate, etc., can be attributed solely to the method of presenting the material.

This kind of control of stimulus materials is not easily achieved when the task involves principle learning. If a concept learning task is chosen, much more control over the materials is possible and greater knowledge of important variables is at hand. For instance, number of examples, ratio and sequence of positive and negative examples, amount and kind of feedback have all been shown to be powerful variables in the concept learning literature. Each can be held constant or systematically varied in presenting a concept by either the expository or the discovery mode.

The following example of a task used by Scott (1970) in a study of discovery vs. expository learning may clarify this recommendation. The concept to be learned was quadrilateral. The relevant attribute was "four sides and only four sides". Size, shape, orientation, and relative lengths of the sides were irrelevant attributes. The task for both discovery and expository groups was to learn to recognize and to produce:

1. examples of the attribute "four sides and only four sides"
2. examples of the concept

3. non-examples of the concept

4. the name of the relevant attribute of the concept

5. the concept definition

Twelve examples were used under both the discovery and expository methods, eight positive examples and four negative examples. The sequence of presentation of positive and negative instances was the same for both methods of presentation, i.e., +, -, +, -, +, +, +, -, +, -, +, +. Exactly the same examples were given under each presentation method. The difference between presentation methods occurred in the statement accompanying each example. Under the expository method, the example was accompanied by a statement such as: "Look at this figure. It has four sides. Count them." Under the discovery method, each example was accompanied by a statement such as: "Look at this figure. How many sides does it have? Count them." Thus, the essential difference between the modes of presentation was the point at which the relevant attributes of the concept were pointed out to S. In the expository group the relevant attributes were given first and then as each example was presented, the relevant attributes were again pointed out. In the discovery group Ss had to induce and hypothesize the relevant attributes as each example was presented. Their hypotheses were not confirmed and the relevant attributes were not named until all examples had been presented. When materials and tasks are specifiable in this manner the differences

between presentation methods are much more easily understood and the findings of various studies more comparable than those now in the literature.

That this is a viable task for studies comparing discovery vs. expository methods of presenting materials is clear from the previous example. Its worth becomes even more clear in relation to its possible use in studies of the effects of amount of guidance. The contention of the authors is that the kind of information given is at least at important as the amount. Several kinds of information might be given to guide the learner in discovery of a concept--the number of relevant attributes, a description of the relevant attributes, examples of the concept, or the concept definition. Both the number of cues and kind of cues given should be noted.

With regard to the study of effects of verbalization on discovery learning or expository learning, this type of task is no less viable. Verbalization may be required for relevant attributes or for concept definitions discovered by or told to Ss. What is verbalized may or may not affect what is learned, but when what is verbalized can be more exactly specified under a paradigm such as this, the findings will be more informative and useful.

Measurement of the Effects of Discovery Learning

In comparing the effects of expository and discovery methods of presentation, it was noted that expository methods were superior when the criterion was short-term retention, while discovery methods were

superior when the criterion was long-term retention or transfer. This suggests that several dependent measures should be used in discovery research.

If the standardized concept learning task described earlier in this paper were utilized, short- and long-term retention of various types of information could be systematically tested by items developed according to a schema proposed by Frayer, Fredrick, and Klausmeier (1969). This schema consists of 12 clearly defined tasks which can be used to test knowledge of a given concept.

Transfer could be tested by using different methods of presenting one concept to two different groups, then presenting a second concept to both groups by one of the methods. Differential performance on the second concept would indicate transfer due to method of presenting the first concept.

Use of the standardized learning task and measures of initial learning, retention, and transfer could lead to more definitive results concerning the relative effects of expository and discovery methods of instruction.

REFERENCES

Anderson, G. L. Quantitative thinking as developed under connectionist and field theories of learning. In E. N. Swenson, et al. (Eds.), Learning theory in school situations. Minneapolis: University of Minnesota Press, 1949. Pp. 40-73.

Bruner, J. S. The act of discovery. Harvard Educational Review, 1961, 31, 21-32.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. A study of thinking. New York: John Wiley, 1956.

Corman, B. R. The effect of varying amounts and kinds of information as guidance in problem solving. Psychological Monographs, 1957, 71 (2, Whole No. 431).

Craig, R. C. The transfer value of guided learning. New York: Teacher's College, Columbia University, 1953.

Craig, R. C. Directed versus independent discovery of established relations. Journal of Educational Psychology, 1956, 47, 223-234.

Cronbach, L. J. The logic of experiments on discovery. In L. S. Shulman and E. R. Keislar (Eds.), Learning by discovery: A critical appraisal. Chicago: Rand McNally, 1966. Pp. 77-92.

Dewey, J. Experience and education. New York: Collier Books, 1963.

Ewert, P. H., & Lambert, J. F. Part II. The effect of verbal instructions upon the formation of a concept. Journal of General Psychology, 1932, 6, 400-413.

Forgus, R. A., & Schwartz, R. J. Efficient retention and transfer as affected by learning method. Journal of Psychology, 1957, 43, 135-139.

Frayer, D. A., Fredrick, W. C., & Klausmeier, H. J. A schema for testing level of concept mastery. Working Paper from the Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, 1969, No. 16.

GPO 822-658-3

Gagné, R. M. The conditions of learning. New York: Holt, Rinehart and Winston, 1965.

Gagné, R. M. Varieties of learning and the concept of discovery. In L. S. Shulman and E. R. Keislar (Eds.), Learning by discovery: A critical appraisal. Chicago: Rand McNally, 1966. Pp. 135-150.

Gagné, R. M., & Brown, L. T. Some factors in the programming of conceptual learning. Journal of Educational Psychology, 1961, 62, 313-321.

Gagné, R. M., & Smith, E. E., Jr. A study of the effects of verbalization on problem solving. Journal of Experimental Psychology, 1962, 63, 12-18.

Glaser, R. Variables in discovery learning. In L. S. Shulman and E. R. Keislar (Eds.), Learning by discovery: A critical appraisal. Chicago: Rand McNally, 1966. Pp. 13-26.

Grote, C. N. A comparison of the relative effectiveness of direct-detailed and directed-discovery methods of teaching selected principles of mechanics in the area of physics. Unpublished doctoral dissertation, University of Illinois, 1960.

Haslerud, G. M., & Meyers, S. The transfer value of given and individually derived principles. Journal of Educational Psychology, 1958, 49, 293-298.

Hendrix, G. A new clue to transfer of training. Elementary School Journal, 1947, 48, 197-208.

Kagan, V. Learning, attention and the issue of discovery. In L. S. Shulman and E. R. Keislar (Eds.), Learning by discovery: A critical appraisal. Chicago: Rand McNally, 1966. Pp. 151-161.

Katona, G. Organizing and memorizing. New York: Columbia University Press, 1940.

Kendler, H. H. Reflections on the conference. In L. S. Shulman and E. R. Keislar (Eds.), Learning by discovery: A critical appraisal. Chicago: Rand McNally, 1966. Pp. 171-176.

Kersh, B. Y. The adequacy of "meaning" as an explanation for superiority of learning by independent discovery. Journal of Educational Psychology, 1958, 49, 282-292.

Kersh, B. Y. The motivating effect of learning by directed discovery. Journal of Educational Psychology, 1962, 53, 65-71.

Kittell, J. E. An experimental study of the effect of external direction during the learning on transfer and retention of principles. Journal of Educational Psychology, 1957, 48, 391-405.

McConnell, T. R. Discovery vs. authoritative identification in children. Studies in Education, 1934, 2 (5), 13-60. Cited in M. C. Wittrock, The learning by discovery hypothesis. In L. S. Shulman and E. R. Keislar (Eds.), Learning by discovery: A critical appraisal. Chicago: Rand McNally, 1966. Pp. 33-76.

Morrisett, L. N. Further reflections. In L. S. Shulman and E. R. Keislar (Eds.), Learning by discovery: A critical appraisal. Chicago: Rand McNally, 1966. Pp. 177-180.

Moss, J.,Jr. An experimental study of the relative effectiveness of the direct-detailed and the directed discovery methods of teaching letterpress imposition. Unpublished doctoral dissertation, University of Illinois, 1960.

Schwartz, B. D. Certain relationships between verbalization and concept formation: The destruction of ideas by words. Unpublished doctoral dissertation, Princeton University, 1948. Cited by B. Worthen, Discovery and expository task presentation in elementary mathematics. Unpublished Master's Thesis, University of Utah, 1965.

Scott, J. A. The effects on short- and long-term retention and on transfer of two methods of presenting selected geometry concepts. Technical Report from the Wisconsin Research and Development Center for Cognitive Learning, The University of Wisconsin, 1970, No. 138.

Stacey, C. L. The law of effect in retained situation with meaningful material. In E. J. Swenson et al., Learning theory in school situations. University of Minnesota Studies in Education. Minneapolis: University of Minnesota Press, 1949. Pp. 74-103.

Swenson, E. J. Organization and generalization as factors in learning, transfer, and retroactive inhibition. In E. J. Swenson, et al., Learning theory in school situations. Minneapolis: University of Minnesota Press, 1949. Pp. 9-39.

Thiele, C. L. The contribution of generalization to the learning of addition facts. New York: Teacher's College, Columbia University, 1938.

Tomlinson, R. M. A comparison of four methods of presentation for teaching technical material. Unpublished doctoral dissertation, University of Illinois, 1962.

Wittrock, M. C. Verbal stimuli in concept formation: Learning by discovery. Journal of Educational Psychology, 1963, 54, 183-190.

Wittrock, M. C. The learning by discovery hypothesis. In L. S. Shulman and E. R. Keislar (Eds.), Learning by discovery: A critical appraisal. Chicago: Rand McNally, 1966. Pp. 33-76.

GPO 822-658-2